**Project Diary**

**Name:** Youcheng ZHANG

**Programme:** Molecular Technique in Life Science

**Course:** Project in molecular Life science (KB8024/KB8025)

|  |  |
| --- | --- |
| **Date** | **Work** |
| Feb 15th, 2018 | 1) Read *Noble WS (2009) A Quick Guide to Organizing Computational Biology Projects. PLOS Computational Biology 5(7): e1000424.* [*https://doi.org/10.1371/journal.pcbi.1000424*](https://doi.org/10.1371/journal.pcbi.1000424)  2) Read [Linux Tips and Tricks](http://bioinfo.se/courses/project-in-molecular-life-science-kb7006/lessons/linux-tips-and-trix/) + [How to organize your project](http://bioinfo.se/courses/project-in-molecular-life-science-kb7006/lessons/organize-project/)  3) Created GitHub account |
| Feb 16th | 1) Practice Bash Command: <http://swcarpentry.github.io/shell-novice/>  2) Practice Git Command: <http://swcarpentry.github.io/git-novice> |
| Feb 17th | 1) Practiced Bash Command, Write bash scripts  2) Start literature search about “Protein secretion”, “Signal peptide”, and “Gram-negative bacteria” |
| Feb 18th | Literature review |
| Feb 19th, 2018 | 1) Reorganized folders on GitHub  2) Finished week1 assignment: createfolder.sh <https://github.com/YouchengZHANG/project/tree/master/assignment/week1> |
| Feb 20th | 1) Add Project Diary  2) Start literature search about “Signal peptide prediction”, “Machine learning approaches”, “Neural network method” and “Random Forest”  3) Start week2 assignment: Summary of 5 relevant papers  4) Finished Python Command: <http://swcarpentry.github.io/python-novice-inflammation/> |
| Feb 21st | 1) Write and Add week2 assignment: List\_of\_papers\_and\_Summary.pdf  <https://github.com/YouchengZHANG/project/tree/master/assignment/week2>  2) Read *Costa, T. R. et al. Secretion systems in Gram-negative bacteria: structural and mechanistic insights. Nature reviews. Microbiology* ***13****, 343-359, doi: 10.1038/nrmicro3456 (2015).*  3) Read *Chou, K. C. Prediction of protein signal sequences. Current protein & peptide science* ***3****, 615-622 (2002).*  4) Write week2 assignment: Project\_plan |
| Feb 22nd | 1) Read *Zeiler, M.D. and Fergus, R., 2014, September. Visualizing and understanding convolutional networks. In European conference on computer vision (pp. 818-833). Springer, Cham*.  2) Learn sklearn and one-hot encoding  3) Extract feature using part of the raw data |
| Feb 23rd | 1) Journal Club and Learn convolutional networks  2) Learn sklearn.svm and cross-validation |
| Feb 24th | 1) Write Feature Extractor  2) Run SVM with different kernel and parameters and Run different cross-validation  3) Write Window size operator  4) Learn how to save and load the trained model |
| Feb 25th | 1) Read and Learn how to process PSSM *Jones, D.T., 1999. Protein secondary structure prediction based on position-specific scoring matrices1. Journal of molecular biology, 292(2), pp.195-202*.  2) Read literatures about “PSSM normalization from raw profile matrix value” and “scaling window size in signal peptide prediction”:  *Sharma, R., Sharma, A. et al. 2015. Predict gram-positive and gram-negative subcellular localization via incorporating evolutionary information and physicochemical features into Chou's general PseAAC. IEEE transactions on nanobioscience, 14(8), pp.915-926.*  *Chou, K.C., 2001. Prediction of signal peptides using scaled window. peptides, 22(12), pp.1973-1979.*  3) Try to write both bash and python scripts for separating each sequence information into single files and then running PSI-BLAST locally  4) Learn additional Bash Command: variable assignment, calculation, input/output, $ sed/awk  5) Learn how to manage background jobs: $ &, $ nohup, $ screen  6) Write PSSM extractor |
| Feb 26th , 2018 | 1) Test PSI-BLAST locally, change parameter e.g. -num\_thread, -word\_size to estimate running time  2) Learn background command: $ jobs, $ ps -aux, $ kill, $ lscpu, $ top -H  2) Write PSSM window size operator  3) Modify the window size parameter from [(i – n)…i…(i + n)] to [(i – m)…i…(i + n)] where the two edges of window size could be different |
| Feb 27th |  |
| Feb 28th |  |
|  |  |
|  |  |
|  |  |